MATHEMATICAL SCIENCES
BELL LABORATORIES
LUCENT TECHNOLOGIES
600 Mountain Ave.
Murray Hill, NJ 07974

Fundamentals of Combinatorial Optimization and Algorithm Design: Final Report

Amount Received: 40K

Duration: 1 Year

Key Personnel:

- Bruce Shepherd, 908-582 4181, bshep@research.bell-labs.com
- Peter Winkler, peter.winkler@dartmouth.edu
- Chandra Chekuri, 908-582 1204, chekuri@research.bell-labs.com

Funded through March 2005.

Authorization: Debasis Mitra, Vice-President

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE May 24, 2005	3. REPORT TYPE AND DATES COVERED **Final Report** - 01 Jan 2005 – 31 Mar 2005

4. TITLE AND SUBTITLE

Fundamentals of Combinatorial Optimization and Algorithm Design

5. FUNDING NUMBERS

N00014-04-M-0042

6. AUTHOR(S)

F. Bruce Shepherd

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lucent Technologies Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

8. PERFORMING ORGANIZATION REPORT NUMBER

052405-03

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Attn: Donald Wagner ONR 311
Ballston Tower One
800 North Quincy Street
Arlington, VA 22217 – 5660

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

12 b. DISTRIBUTION CODE

A

13. ABSTRACT (Maximum 200 words)

The main activities supported under this grant are research and support for C. Chekuri, B. Shepherd and P. Winkler. Funds also supported one summer intern, Andrew McGregor from U Penn, who worked with Shepherd on recognizing Hilbert Bases and other theoretical topics in Math Programming. Visits from scientists include a 2-week visit from Gianpaolo Oriolo (Rome), which resulted in new joint work on robust network design, a week visit from Seffi Naor (Technician) and a visit from Anreas Sebo who spoke on a new result of Bessy and Thomasse that solves an old conjecture of Gallai.. Research highlights this year include: proof that in planar graphs with all capacities at least 2, the intergrality gap for edge-disjoint paths is polylogarithmic (the paper was invited for the selected papers issue devoted to FOCS 2005); a first result showing the hardness of the robust network design and introduction of the single-source hose model for robust networks; and an unlikely question: is it hard to determine whether the rows of 0,1 matrix form a Hilbert Basis? Conferences attended were the 2004 APPROX/RANDOM (Chekuri), CORC 4th Optimization Day, INOC, Aussois workshop (Shepherd), and .a workshop in Bertinoro, Italy (Chekuri & Shepherd).

14. SUBJECT TERMS	15. NUMBER OF PAGES **1**
	16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	**UNCLASSIFIED**	**UNCLASSIFIED**	**UNLIMITED**

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Introduction

The main activities being supported under this grant, are research support (primarily for travel) for C. Chekuri, B. Shepherd, and P. Winkler. Second to this, funds were used to support one summer intern, and several brief visits from scientists. Peter Winkler has now gone (as of July 2004) on leave to a position at Dartmouth University, so he will no longer request further support from this grant. Part of his travel support will subsidize Shepherd's attendance at a meeting in Aussois and then to attend the International Network Optimization Conference (INOC).

Research highlights this year include:

- In [5] it was proved that in planar graphs with all capacities at least 2, the integrality gap for (the usual formulation for) edge-disjoint paths is polylogarithmic; this is an exponential decrease from the $\sqrt{n}$ gap for planar graphs with arbitrary capacities. (The paper was invited for the selected papers issue devoted to FOCS 2005.)

- A first result showing the hardness of the robust network design and introduction of the single-source hose model for robust networks.

- An unlikely question: is it hard to determine whether the rows of a $0, 1$ matrix form a Hilbert Basis?

We hosted 1-2 week (Summer) visits from Anupam Gupta (CMU) and Adrian Vetta (McGill). We also had a week visit from Seffi Naor (Technion) and a subsequent 2-day follow up meeting to work with C. Chekuri in February. We hosted a summer intern Andrew McGregor from UPenn (under the supervision of Sampath Kannan) who worked with Shepherd on recognizing Hilbert Bases and other theoretical topics in Math Programming. We also supported a 2-week visit by Gianpaolo Oriolo (Rome) which resulted in some new joint work on robust network design. Finally, Andras Sebö made a short visit and spoke on a new result of Bessy and Thomassé that solves an old conjecture of Gallai.

Travel support consisted of funding Chekuri and Shepherd to attend (last May) a interesting workshop in Bertinoro, Italy. This was organized by Leen Stougie, Michel Goemans and Martin Skutella and brought together researchers in CS, OR and combinatorics. Chekuri spoke on work of himself and Shepherd on integer decomposition results for Steiner Forests and related network design problems [6]. In addition, Shepherd attended the International Network Optimization Conference (INOC) and the Aussois workshop on integer programming and combinatorial optimization.

We summarize below the work completed and in progress. Please contact Bruce Shepherd (bshep@lucent.com) as we are very happy to provide any of the work in progress as manuscripts.

Integrality Gaps and Approximation Algorithms for Maximum Throughput Problems

In a throughput problem we are given a graph G and terminals $s_i, t_i, i = 1, 2 \ldots, k$, and we wish to find a maximum *routable subset* of $\{1, 2, ..., k\}$. The notion of "routability" depends on the application at hand. For instance, in the edge-disjoint paths problem (EDP) a subset S is routable, if there is a collection of edge-disjoint paths P_i : $i \in S$, such that each P_i joins s_i, t_i. EDP is one of the most fundamental problems in combinatorial optimization. Apart from its applications in VLSI design and network design, it is also intrinsic to many approaches for solving other applied problems such as scheduling. For instance, Shepherd and Matthew Andrews created a scheduling system within Lucent, based on EDP.

In [9] it was shown that in directed graphs it is hard to approximate (in polytime, if P not equal to NP) the optimum to within a factor of $m^{.5-\epsilon}$ for any $\epsilon > 0$, and hence EDP is terribly difficult with respect to polytime approximations. Their reduction however breaks down if small edge congestions are allowed, and also breaks down for undirected graphs. In fact, the gap for undirected graphs is only known to lie in the range 2 to $m^{.5-\epsilon}$.

In [4], we show the first positive results in this direction by showing that the all-or-nothing multicommodity flow problem problem does admit a poly-logarithmic (i.e., polynomial in the variable $\log(n)$) approximation for general graphs – please refer to the paper: `http://cm.bell-labs.com/cm/ms/who/bshep/pub.html`. (The all-or-nothing flow problem is the throughput problem where a subset is routable, if the $s_i - t_i$ pairs admit a multicommodity flow. Note that unlike EDP, determining whether a subset is routable is easy for this problem, thus the hard part is the subset selection aspect.) The techniques in this paper used recent results of Räcke on oblivious routing as well as some interesting graph theoretic clustering results in 2-connected graphs.

Ideas from the work on all-or-nothing flow inspired the authors to take a look at EDP itself. Using schemes of Robertson, Seymour and Thomas, the authors have since proved the following theorem in planar graphs. We call a set X *well-linked* in G, if for any subset S with $|S \cap X| \leq |X - S|$, we have $|\delta(S)| \geq |S \cap X|$. We show that if S is well-linked in a planar graph G, there is a constant C (about $10,000$ at present unfortunately) such that for any matching M of size $|S|/C$ on S, we can find paths connecting the endpoints of M, such that each edge lies in at most 2 of them. This result implies (again using Räcke's results) that EDP can be approximated to within a polylogarithmic factor where only $\sqrt{n}$ was previously best known. Alternatively, this can be viewed as saying that the natural LP formulation for EDP has a polylogarithmic integrality gap for planar graphs with all capacities at least 2. In contrast, if some capacities are at most 1, then this gap may be $\Omega(\sqrt{n})$, exponentially larger. Thus, one goal for further work is to find cutting planes that tightens the integrality gap for arbitrary capacities. We mention that recently J. Kleinberg has used our framework to strengthen the planar result to the case of Eulerian planar graphs (not just all capacities at least 2). The techniques used, also show that there is a constant approximation for product multicommodity flow in planar graphs, thus giving a new proof of an earlier result of Klein, Plotkin and Rao for uniform multicommodity flow in planar graphs.

This work appears in [5].

Continuing work has recently involved exploring decomposition methods for high-degree constant expansion graphs. Some progress has been made in particular on shortening of the proofs of Räcke's celebrated result (however this thread is still under development!) Most recently, the authors have found a direct decomposition into well-linked sets that avoids the use of Räcke's decomposition. One interesting side-effect is that it allows the authors to obtain similar polylog-approximation algorithms for the node-disjoint versions of the problem. This is not possible by Räcke's decomposition since in that case $\Omega(\sqrt{n})$ gaps are known in the node versions of oblivious routing [10]. This work is appearing in STOC 2005 in the paper *Multicommodity flow, well-linked terminals, and routing problems* [8].

Robust Network Optimization

During Gianpaolo Oriolo's visit to Bell Labs, the authors explored the problem of robust network design.

Network designers have traditionally adopted the view that an accurate estimate for point-to-point traffic is given a priori. With the increasing importance of flexible services (such as VPNs or remote storage/computing), there has been increasing interest in the design of networks for situations where traffic patterns are either not well known a priori or changing rapidly. In these settings the network should be "dimensioned" (i.e., assigned capacity) to support not just one traffic matrix, but a larger class of matrices determined by the application. This results in a *robust optimization* problem, where we are given a *universe* $\mathcal{U}$ of demand matrices (normally specified as a convex region), and the goal is to design a minimum cost network so that every demand matrix in $\mathcal{U}$ can be *supported.* The simplest form of this problem where fractional capacities are allowed was introduced by Ben-Ameur and Kerivin [1], but only recently was it shown to be NP-hard [7] by Chekuri, Oriolo, Scutella and Shepherd. The problem considered is how to allocate fractional link capacities that are sufficient to support every demand, i.e., so that there is a multicommodity flow for every demand matrix in the universe $\mathcal{U}$. It is perhaps the fact that capacities are allowed to be fractional that makes this result somewhat surprising.

The main complication in proving hardness is the fact that there is no max-flow min-cut theorem for multicommodity flow. To overcome this, the authors introduce a "trick" of considering a more simply analyzed class $\mathcal{U}$ of matrices they call *single-source hose demand matrices*. This single source problem is defined as follows. There is a given root node r with a specified marginal traffic value b_r, and each other node also has a marginal value b_v. The authors normally consider that b_v's are 0, 1 and that b_r is some integer. A matrix d is then a *single-source hose demand* (i.e., it will be in our universe $\mathcal{U}$) if $d_{ij} = 0$ unless $i = r$ and $\sum_j d_{rj} \leq b_r$ and each $d_{ri} \leq b_i$. We then ask for the minimum fractional capacity so that every such single-source demand matrix can be routed.

There are some special cases of this single-source robust design problem of interest. In the case where all b_v's are 1 and b_r is also, then the optimal fractional capacity

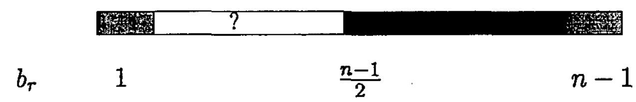

Figure 1: Known complexity status for the single-source robust network design problem.

allocation is obviously just a minimum spanning tree. If only some of the b_v's are 1, then it is the fractional relaxation for the undirected Steiner tree problem. In the case, where all b_v's are 1, and $b_r = n$, then the optimal solution is a shortest path tree rooted at r. So at the extreme values of $b_r = 1$ or n, the problem is well-known to be polytime computable.

The authors show that the problem becomes hard for values of $b_r = cn$ where $c \in [1/2, 1)$, and suspect this continues to hold for every $c \in (0, 1)$ – see Figure . In this case, [7] gives a reduction from checking whether a graph has the expander property. In fact, based on results in [3] the authors show that (assuming a certain conjecture in complexity theory), if $b_r = n/2$, then for any $\epsilon > 0$ it is hard to approximate the network to within a factor of $2 - \epsilon$. Moreover, they exhibit a matching 2-approximation algorithm in this special case.

Hilbert Bases and Dual Integrality

While McGregor had previously primarily worked in coding theory and approximation algorithm, he decided to spend the summer learning more about the basic theory of math programming. Supplied with the book of Schrijver and Cornuéjols, he has learned most of the basic definitions (face, cones, Hilbert bases, branch and bound) quickly. McGregor and Shepherd worked on one of the 10 celebrated conjectures in Cornujols book: that in order to check total dual integrality of an ideal matrix, it is enough to check each $0, 1, \infty$ valued objective. Preempted by subproblems in this effort, the two spent most of their time examining the integrality and toal dual integrality of problems of the form: $P^*(A) = \{x : Ax \leq 1\}$, for $0, 1$ A. While the theory of such problems is well known in the presence of nonnegativity constraints, there does not appear (yet!) to be literature on the version where negativity is allowed. We refer to a matrix as P^* if $P^*(A)$ is integral, and *TDI-P^** if the corresponding system is TDI. We have noted that P* matrices include balanced and $0, 1$ TUM matrices, and are not included in the class of perfect matrices. We also explored a conjecture about when a matrix is TDI-P* based on whether the rows of A form a Hilbert basis. Finally, we asked what seems a very natural question, which to our knowledge (which includes some probing of colleagues) has not been considered. What is the complexity status of determining whether the rows of a $0, 1$ matrix form a Hilbert basis. At first glance, this would seem to be certainly NP-hard (although in fixed dimension, it was shown to be in P by R. Kannan). We tried to build up some operations that may be useful in attacking this problem. We are starting to wonder whether this might not be a polytime recognizable class!

Workshop/Conference Travel

Chekuri:

1. Workshop on Combinatorial Optimization (organizers Michel Goemans, Martin Skutella and Leen Stougie), May 1-6, 2004, Bertinoro, Italy
2. STOC 2004, Chicago, IL June 13-15, 2004
3. APPROX 2004: August, 2004, Cambridge, MA.

Shepherd:

1. Workshop on Combinatorial Optimization (organizers Michel Goemans, Martin Skutella and Leen Stougie), May 1-6, 2004, Bertinoro, Italy
2. SODA 2005: 16th Annual ACM-SIAM Symposium on Discrete Algorithms, January 23-25, 2005, Vancouver, British Columbia, Canada.
3. Aussois Workshop on Combinatorial Optimization and Integer Programming, March 13-20, 2005, Aussois, France
4. International Network Optimization Conference (INOC 2005), March 20-23, 2005, Lisbon, Portugal.

Winkler:

1. Cornell Markov Chain Workshop, Ithaca, NY, May 9, 2004
2. DIMACS biogame conference, New Brunswick, NJ April 5, 2004
3. Columbia Theory Day, May 2004
4. Various visit to Institute for Advanced Studies, Princeton, NJ

References

[1] W. Ben-Ameur, H. Kerivin, Routing of uncertain demands, *Optimization and Engineering.*

[2] S. Bessy, S. Thomassé, Spanning a strong digraph by α circuits: A proof of Gallai's conjecture, to appear in *Combinatorica.*

[3] S. Chawla, R. Krauthgamer, Ravi Kumar, Y. Rabani, and D. Sivakumar. On the Hardness of Approximating Multicut and Sparsest-Cut. Submitted, November 2004.

[4] C. Chekuri, S. Khanna, and F.B. Shepherd, The All-or-nothing multicommodity flow problem, *Proceedings of the 36th ACM Symposium on Theory of Computing*, 2004.

[5] C. Chekuri, S. Khanna, and F.B. Shepherd, Edge-disjoint paths in planar undirected graphs, *FOCS, 2005.*

[6] C. Chekuri and F.B. Shepherd, Splitting-off for Approximate Integer Decomposition of Network Design Problems. Submitted.

[7] C. Chekuri, G. Oriolo, M.G. Scutella, and F.B. Shepherd. Hardness of Robust Network Design, *INOC 2005.*

[8] C. Chekuri, S. Khanna, and F.B. Shepherd, Multicommodity flow, well-linked terminals, and routing problems, *STOC 2005.*

[9] Venkat Guruswami, Sanjeev Khanna, Raj Rajaraman, Bruce Shepherd, Mihalis Yannakakis. Near-Optimal Hardness Results and Approximation Algorithms for Edge-Disjoint Paths and Related Problems, *Journal of Computer and System Sciences*, Volume 67, Issue 3, Nov. 2003, Pages 473-652. (also in *Proc. 31st Annual ACM Symposium on Theory of Computing (STOC)*, pp. 19-28, 1999)

[10] Mohammad Hajiaghayi, Robert Kleinberg, Tom Leighton, and Harald Racke. Oblivious Routing in Node-Capacitated and Directed Graphs. In Proceedings of the *16th ACM-SIAM Symposium on Discrete Algorithms (SODA 2005)*, pages 782-790.